

DNA Barcoding of genus *Hexacentrus* in China reveals cryptic diversity within *Hexacentrus japonicus* (Orthoptera, Tettigoniidae)

Hui-Fang Guo¹, Bei Guan¹, Fu-Ming Shi¹, Zhi-Jun Zhou¹

¹ The Key Laboratory of Zoological Systematics and Application, College of Life Sciences, Hebei University, Baoding 071002, China

Corresponding author: Zhi-Jun Zhou (zhijunzhou@163.com)

Academic editor: F. Montealegre-Z | Received 31 March 2016 | Accepted 25 May 2016 | Published 7 June 2016

<http://zoobank.org/C6C19258-1B3C-4CDC-B334-EA6F2C5B09A3>

Citation: Guo H-F, Guan B, Shi F-M, Zhou Z-J (2016) DNA Barcoding of genus *Hexacentrus* in China reveals cryptic diversity within *Hexacentrus japonicus* (Orthoptera, Tettigoniidae). ZooKeys 596: 53–63. doi: 10.3897/zookeys.596.8669

Abstract

DNA barcoding has been proved successful to provide resolution beyond the boundaries of morphological information. Hence, a study was undertaken to establish DNA barcodes for all morphologically determined *Hexacentrus* species in China collections. In total, 83 specimens of five *Hexacentrus* species were barcoded using standard mitochondrial cytochrome c oxidase subunit I (COI) gene. Except for *H. japonicus*, barcode gaps were present in the remaining *Hexacentrus* species. Taxon ID tree generated seven BOLD's barcode index numbers (BINs), four of which were in agreement with the morphological species. For *H. japonicus*, the maximum intraspecific divergence (4.43%) produced a minimal overlap (0.64%), and 19 specimens were divided into three different BINs. There may be cryptic species within the current *H. japonicus*. This study adds to a growing body of DNA barcodes that have become available for katydids, and shows that a DNA barcoding approach enables the identification of known *Hexacentrus* species with a very high resolution.

Keywords

BOLD, China, DNA Barcoding, *Hexacentrus*, species delineation

Introduction

DNA barcoding employs short, standardized gene regions (5' segment of mitochondrial cytochrome oxidase subunit I for animals) as an internal tag to enable metazoan species identification (Hebert et al. 2003). Schmidt et al. (2015) found that DNA barcoding largely supported 250 years of classical taxonomy for central European bees. Unlike distinct species, closely related species offer a great challenge for phylogeny reconstruction and species identification with DNA barcoding due to overlapping genetic variation (Dai et al. 2012). For example, Versteirt et al. (2015) found that DNA barcoding offered a reliable framework for mosquito species identification in Belgium except for some closely related species. Zhou et al. (2012) found that molecular identification with DNA barcoding supported most traditional morphological species of genus *Ruspolia* in China.

In this study, our objective is to assess the utility of DNA barcoding for closely related katydid species, belonging to the genus *Hexacentrus* (Serville, 1831) in China. *Hexacentrus* is mainly distributed in Australian, Afrotropical and Oriental realms. *Hexacentrus* is a particularly speciose genus, containing 24 known species (Eades et al. 2016). *Hexacentrus* was the single genus within Hexacentrinae, which has been reported in China according to “Orthoptera Species File” (Eades et al. 2016). Up to now, a total of six *Hexacentrus* species have been reported, including *H. japonicus* (Karny, 1907), *H. unicolor* (Serville, 1831), *H. yunnaneus* (Bey-Bienko, 1962), *H. fuscipes* (Matsumura & Shiraki, 1908), *H. mundus* (Walker, 1869) and *H. expansus* (Wang and Shi 2005). Due to the rather difficult morphological discrimination between *Hexacentrus* species, an interspecific molecular delineation was needed. To make *Hexacentrus* more accessible to the scientific community, the open access project “BHC” had been initiated in the Barcoding of Life Data systems (BOLD) (Ratnasingham and Hebert 2007). There are a few DNA barcoding study concentrated on *Hexacentrus*, and the number of barcode sequences was limited in BOLD. The goals of this study are as following: (i) it will allow scientists with molecular capability but insufficient knowledge of *Hexacentrus* taxonomy and systematics to recognize species and document the biodiversity of *Hexacentrus*. (ii) For Tettigoniidae taxonomists, it contributes to integrative taxonomic approaches, such as the elucidation of related species and clarification of problematic species groups, association of the sexes within one species, and the identification of new species (Gibbs 2009, 2011, Packer et al. 2009, Schmidt et al. 2015). To this end, we checked for the presence of species barcode gaps and cryptic diversity within species. BOLD’s barcode index number (BIN) analysis tool (Ratnasingham and Hebert 2013) was used to analyze *Hexacentrus*.

Material and methods

Collection of specimens

All specimens were collected by hand or sweeping method during their active season (July–November). *Hexacentrus* species were all gathered in China from 12 localities, with

the latitude from 18.70°N to 41.80°N and the longitude from 97.83°E to 123.38°E. One or more specimens were chosen from each locality in order to include as many morphologically distinguishable individuals per site as possible. Specimens were collected and stored in 100% ethanol at -20 °C and were deposited in the Hebei University Museum. Species-level identification was based on the original morphological descriptions, locality data and additional information. Details on all specimens (sampling location, GPS coordinates, voucher number, BOLD number, etc.) are available within the “DNA Barcoding of *Hexacentrus* in China, BHC” project in the Barcode of Life Data Systems (BOLD. www.barcodinglife.org).

DNA extraction, amplification and sequencing

Total DNA was extracted from the muscle of one hind leg of each specimens using TI-ANamp Genomic DNA Kit in accordance with the manufacturer's instructions. The standardized gene regions of animals DNA barcoding was amplified using the primers COBU (5'-TYT CAA CAA AYC AYA ARG ATA TTG G-3') and COBL (5'-TAA ACT TCW GGR TGW CCA AAR AAT CA-3') (Pan et al. 2006). The 50 µL polymerase chain reaction (PCR) mixture contained 3 µL of template DNA, 5 µL of 10 × buffer, 4 µL of dNTP mix, 5 µL of each primer (10 µM each), 0.5 µL of *Taq* polymerase (5 U/µL), and 27.5 µL of water. The thermal profile was: 94 °C for 3 min, 34 cycles at 94 °C for 30 s, 49 °C for 30 s, and 72 °C for 90 s, and final extension at 72 °C for 8 min. PCR products were visualized in 1% agarose gels electrophoresis. PCR products were sequenced directly using ABI BigDye Terminator chemistry on ABI3730 automated sequencer (Applied Biosystems) in Genewiz Inc. (Beijing, China), and in both directions to minimize PCR artifacts, ambiguities and base calling error.

Data analysis

Consensus sequence of both directions was assembled using SeqMan in Lasergene and verification of ambiguities and unexpected stop codons were performed in EditSeq (Burland 2000). Sequence alignments were conducted using Clustal X 1.81 (Thompson et al. 1997) with default parameters. The both ends of the sequences matching the primer sequences were excised to remove artificial nucleotide similarity introduced by PCR amplification, resulting in the final data sets for barcoding analysis.

The analyses were restricted to the subset of sequences, which met barcode standards (sequence length > 500bp, < 1% ambiguous bases, bidirectional sequencing, country specification). Intra- and inter-specific genetic distances were based on the Kimura-2-parameter (K2P) model (Kimura 1980) using the ‘distance summary’ tool in BOLD. The barcode gap was defined by intraspecific vs. interspecific [nearest neighbor (NN)] genetic distance of species. A globally unique identifier (i.e. BIN) then was assigned to each sequence cluster, creating an interim taxonomic system because the

members of a particular BIN often correspond to a biological species. Character based DNA barcoding used the nucleotide variation in each position across DNA regions as diagnostic characters.

Results

COI sequences were recovered from 86 of the 91 specimens that were analyzed with barcode compliant records from 83 specimens representing five species. Three records have no barcode compliant records because of low quality of trace file. A number of 80 barcodes belong to four previously identified species whereas three analyzed specimens were only identified to genus level because they are female; they probably are *H. yunnanensis* due to collection in Yunnan and separated from other specimens. The 658bp length sequences without indel (insertion/deletion) had full-length records. COI sequences were translated to amino acid sequences to check for stop codons and shifts in reading frame that might indicate the presence of nuclear mitochondrial copies (numts), but none were detected. Diagnostic character analysis was consistent with that of traditional external appearance discrimination. *H. expansus* only having one specimen was not analyzed, thus lacking diagnostic character.

Distance summary and Barcode Gap analysis

Mean intraspecies divergence was 1.32% (ranged between 0.57% and 2.43%), and maximum intraspecies divergence 4.43% was observed in *H. japonicus* (Table 1). When correcting for the uneven sample sizes of species, the within-species divergence decreased from 1.32% to 1.23%. Between *Hexacentrus* species, the average K2P genetic distance was 12.54%, whereas minimum genetic distance only 3.79% (Table 1).

Singleton species (*H. expansus*) were excluded from barcode gap analysis. Except for *H. japonicus*, barcode gap was present in the remaining *Hexacentrus* species (Fig. 1). Although the maximum intraspecific divergence of *H. unicolor* was more than 2% (ranged between 0 and 3.79%), but still less than minimum interspecific between *H. unicolor* and its NN (Nearest Neighbor) *Hexacentrus* spp. However, the maximum intraspecific divergence of *H. japonicus* (4.43%) produced a minimal overlap (0.64%).

Table 1. Mean and maximum intraspecific and nearest neighbor (NN) distance for all specimens.

Species	Mean intraspecific distance	Max intraspecific distance	Nearest neighbor	Distance to NN
<i>Hexacentrus expansus</i>	N/A	N/A	<i>Hexacentrus unicolor</i>	13.19
<i>Hexacentrus japonicus</i>	2.43	4.43	<i>Hexacentrus mundus</i>	3.79
<i>Hexacentrus mundus</i>	0.57	0.93	<i>Hexacentrus japonicus</i>	3.79
<i>Hexacentrus</i> sp.	0.72	1.08	<i>Hexacentrus unicolor</i>	9.72
<i>Hexacentrus unicolor</i>	1.19	3.79	<i>Hexacentrus</i> sp.	9.72

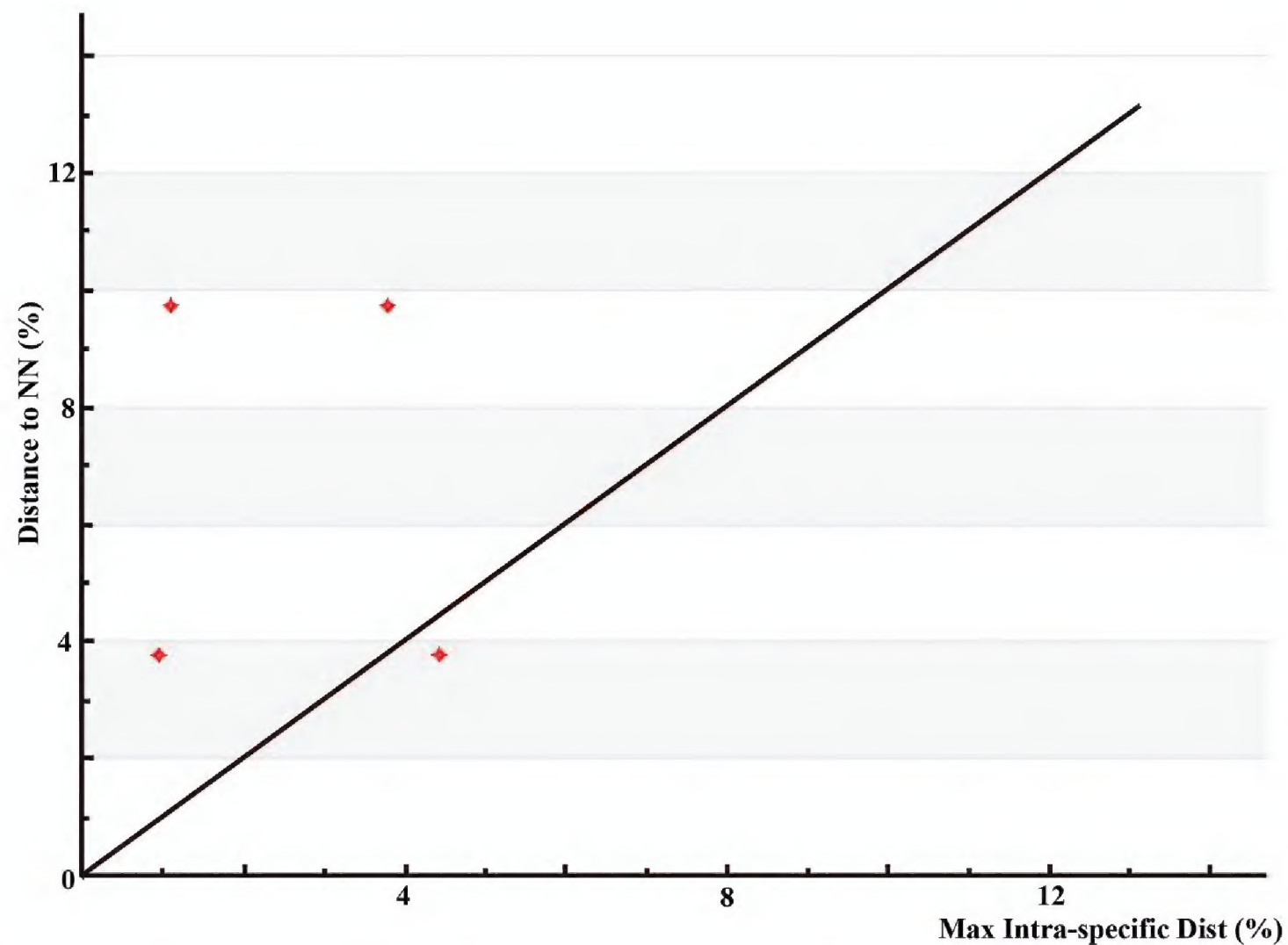


Figure 1. Barcode gap plot showed the distance to the nearest neighbor (NN) vs. the maximum intraspecific distance Kimura-2-parameter (K2P) for 83 specimens. Dots above the 1:1 line indicated the presence of a barcode gap.

Taxon ID tree analysis

The taxon ID tree was divided into seven clades represented by different BINs (Fig. 2). *Hexacentrus unicolor*, *H. mundus* and *H. expansus* were composed well-supported monophyletic groups, which were fully congruent with the morphological species. *Hexacentrus japonicus* were divided into three different BINs. Although no morphological differences were observed among these three BINs, there might be cryptic species within the current *H. japonicus*.

Diagnostic characters analysis

Forty-four diagnostic characters were found in the study (Table 2). Four *Hexacentrus* species had diagnostic characters and the success rate was 80%. The number of diagnostic character sites of *H. spp.* was no less than 20 (Table 2), which may be caused by scarce specimens and by the relative distance of the phylogenetic relationship to others.

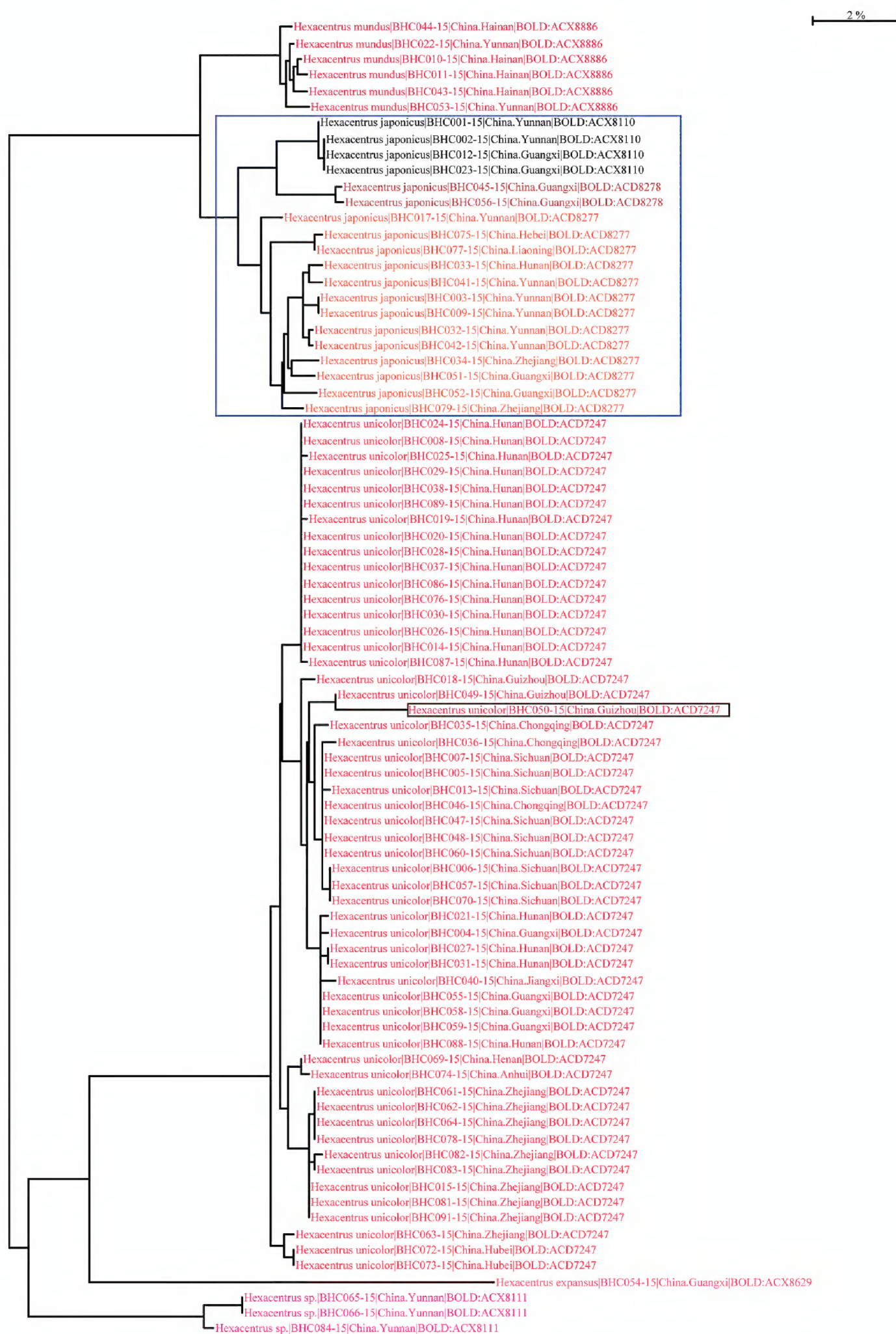


Figure 2. Taxon ID tree revealed seven well-differentiated haplogroups. Process ID, location, and BINs were shown in the tree. The clusters with a blue box indicated there may be two new putative ‘cryptic species’ within *H. japonicus*. The clade with a black box indicated the specimen had more mutation within *H. unicolor*.

Table 2. Character-based DNA barcodes for four *Hexacentrus* species of COI gene. A_7* means A is at the 7th position.

Species	Diagnostic characters	Characters no.	Specimen no.
<i>Hexacentrus</i> sp.	A_7* G_184 C_247 A_301 A_304 T_346 G_391 C_400 G_424 C_463 C_472 C_500 T_517 G_550 G_586 C_607 A_619 C_622 C_625 G_628	20	3
<i>Hexacentrus unicolor</i>	T_25 T_136 T_223 A_227 T_322 T_379 T_424 G_487 A_502 C_517 T_529 C_530 G_532 C_550 C_586 G_619 T_631	17	54
<i>Hexacentrus mundus</i>	T_34 T_118 C_187 T_397 A_643	5	6
<i>Hexacentrus japonicus</i>	T_460 C_514	2	19

Discussion

The present study evaluated the efficacy of using DNA barcodes for the identification of *Hexacentrus* in China and provided a group of sequences associated with the identified species. Using these DNA barcoding, not only can one delineate the boundaries between species, but also assign taxonomic status to unknown specimens from known species.

Hexacentrus unicolor was controversial, and *H. plantaris* (Burmeister, 1838) and *Tedla sellata* (Walker, 1869) were considered as its synonyms. *H. unicolor* is distributed in south of the Yangtze River. In this study, however, one specimen was collected from Henan. In fact, molecular data support *H. unicolor* as a single group, with all specimens sharing one BIN (BOLD: ACD 7247). The specimen with black box (Fig. 2) came from Guizhou, and had more mutations compared to the other *H. unicolor* specimens. We cannot be completely certain that this phenomenon was due to geographic isolation rather than sequencing or calibrating errors. *Hexacentrus japonicus* was closely related to *H. unicolor*, and they are widely distributed in the southwest, central and south areas of China. The tegmina of male *H. japonicas* was short and broad (about 2.75–3.00 times), whereas *H. unicolor* was long and narrow (about 2.95–3.30 times) as long as broad (Wang and Shi 2005). Nevertheless, the analyses revealed that *H. japonicus* ranged to Hebei and Liaoning. Interestingly, *H. japonicus* contained three BINs (BOLD: ACX 8110, BOLD: ACD 8277, BOLD: ACD 8278) with high intraspecific distance (> 2%). This group included 20 specimens, in which no morphological differences were found. Therefore it is necessary to clarify the status of this species complex as it may include more species than currently recognized. *Hexacentrus mundus* was only recorded in Guangxi and Yunnan. In this study, four specimens from Hainan were also identified as *H. mundus* because only 2–3 larger teeth in the middle part of stridulatory file, however there are 6–7 large teeth in *H. japonicas* and *H. unicolor* (Wang and Shi 2005). Thus *H. mundus*’s distribution was enlarged. All specimens of *H. mundus* were assigned as 1 BIN, which clearly confirm a consistency between molecular and morphological analyses. *Hexacentrus expansus*, due to the obviously inflated male tegmina, was easy to identify only by the morphological method. BIN assignments also revealed that *H. expansus* was a separated clade with only a single male available for analysis. The specimens from Yunnan almost certainly represent the ‘true’ *H. yunnaneus* because the type loca-

tion of this species (Hekou, Yunnan) was in close proximity. Hence, further specimens are needed to be analyzed, especially male material.

DNA barcoding, as one effective tool in insect taxonomy, had been already applied widely. It can rapidly acquire molecular data, simplifying species classification and identification. Yet, DNA barcodes has been argued to be unreliable for consistent species identification by many authors due to a number of drawbacks (DeSalle et al. 2005; Will et al. 2005; Rubinoff et al. 2006; Ebach 2011). Recent speciation, incomplete lineage sorting, interspecific hybridization and infection by endosymbiotic bacteria such as *Wolbachia* (Funk and Omland 2003) may all interfere with the performance of DNA barcoding in insects (Virgilio et al. 2010). In this context, most of the species can be amplified successfully; however, five specimens cannot be translated and three specimens only had one represented trace file, which cannot meet the DNA barcoding standards. The deep inspection of trace files indicated that most of these failures arose from co-amplification of the bacterial endosymbiont *Wolbachia*, which disturbed normal interpretation of trace file. It had been estimated that *Wolbachia* is present in two-thirds of all insect species (Hilgenboecker et al. 2008). There was no reason to doubt the absence of *Wolbachia* in *Hexacentrus*. Zhou et al. (2014) found that the nuclear sequence of mitochondrial (numts) reported in *Mecopoda niponensis* may form a separate clade. The same case was reported in *Podisma pedestris* (Bensasson et al. 2000). But in this study, the clades in *H. japonicus* were not caused by numts for all sequences analyzed without translation early termination, base indel, frameshift mutations. On the other hand, geographic isolation was also rejected, for only the specimens from Yunnan included three BINs in this group. The most probable reason was the existence of cryptic species compared to other *Hexacentrus* species.

Species boundaries were hard to delimit only based on morphology, and analyses including additional sources of information such as molecular data, biogeography, behavior and ecology has been called integrative taxonomy which has been shown to be very useful (Dayrat 2005). We are convinced that DNA barcoding can promote the *Hexacentrus* species identification. Our study showed that all known *Hexacentrus* species could be delimiting rapidly through DNA barcoding in China, except for *H. japonicus*. *Hexacentrus japonicus* was problematic when using BOLD analysis. Finally, regardless of the promising results, the incorporation of nuclear genes is valuable for species delimitation and might strengthen the results, as they are independent of the maternal inherited mitochondrial genes.

The ideal situation would be that each species was represented by sequence from its type material, particularly the holotype. Type specimens were also dried specimens and DNA degraded at different level, so not only amplification was difficult, but also the damage of specimens can't be neglected. Recently, Prosser et al. (2016) successfully obtained sequences from century-old type specimens using next-generation sequencing (NGS). We believed that DNA barcoding is useful in revealing cryptic biodiversity, potentially facilitating traditional taxonomy in future.

Acknowledgements

We are grateful to the following people for collecting specimens and providing valuable comments during the manuscript preparation: Qiong Song; Ji-Yuan Feng; Zhi-Lin Chen; Bao-Jie Du; Xun Bian. This study was supported by National Natural Science Foundation of China (No. 31471985), and Excellent Youth Scholars Program of Higher Education of Hebei Province (No. BJ2014006).

References

- Bensasson D, Zhang DX, Hewitt GM (2000) Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Molecular Biology and Evolution* 17: 406–415. doi: 10.1093/oxfordjournals.molbev.a026320
- Bey-Bienko GY (1962) Results of the Chinese-Soviet zoological-botanical expeditions to south-western China 1955–1957. New or less known Tettigonioidea (Orthoptera) from Szechuan and Yunnan. *Trudy Zoologicheskogo Instituta, Leningrad* 30: 135–136.
- Burland TG (2000) DNASTAR's Lasergene sequence analysis software. *Method in Molecular Biology* (Clifton, N.J.) 132: 71–91.
- Burmeister H (1838) Kaukerfe, Gymnognatha (Erste Hälfte: Vulgo Orthoptera). *Handbuch der Entomologie* 2: 235–756.
- Dai QY, Gao Q, Wu CS, Chesters D, Zhu CD, Zhang AB (2012) Phylogenetic reconstruction and DNA barcoding for closely related pine moth species (*Dendrolimus*) in China with multiple gene markers. *PLoS ONE* 7: e32544. doi: 10.1371/journal.pone.0032544
- Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85: 407–415. doi: 10.1111/j.1095-8312.2005.00503.x
- DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360: 1905–1916. doi: 10.1098/rstb.2005.1722
- Eades DC, Otte D, Cigliano M, Braun H (2016) Orthoptera species file. Version 5.0/5.0.
- Ebach MC (2011) Taxonomy and the DNA barcoding enterprise. *Zootaxa* 2742: 67–68.
- Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 34: 397–423. doi: 10.1146/annurev.ecolsys.34.011802.132421
- Gibbs J (2009) Integrative taxonomy identifies new (and old) species in the *Lasioglossum* (*Dialictus*) *tegulare* (Robertson) species group (Hymenoptera, Halictidae). *Zootaxa* 2032: 1–38.
- Gibbs J (2011) Revision of the metallic *Lasioglossum* (*Dialictus*) of eastern North America (Hymenoptera: Halictidae: Halictini). *Zootaxa* 3073: 1–216.
- Hebert PD, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B* 270: S96–99. doi:10.1098/rsbl.2003.0025

- Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008) How many species are infected with *Wolbachia*? – a statistical analysis of current data. *FEMS Microbiology Letters* 281: 215–220. doi: 10.1111/j.1574-6968.2008.01110.x
- Karny HH (1907) *Revisio Conocephalidarum*. Abhandlungen der KK Zoologisch-botanischen Gesellschaft in Wien 4: 98–114.
- Matsumura S, Shiraki T (1908) Locustiden Japans. The journal of the College of Agriculture, Tohoku Imperial University 3: 1–80.
- Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology* 3: e422. doi: 10.1371/journal.pbio.0030422
- Packer L, Gibbs J, Sheffield C, Hanner R (2009) DNA barcoding and the mediocrity of morphology. *Molecular Ecology Resources* 9: 42–50. doi: 10.1111/j.1755-0998.2009.02631.x
- Pan C, Hu J, Zhang X, Huang Y (2006) The DNA barcoding application of mtDNA COI gene in seven species of Catantopidae (Orthoptera). *Entomotaxonomia* 28: 103–110.
- Ratnasingham S, Hebert PD (2007) BOLD: the barcode of life data system. *Molecular Ecology Notes* 7: 355–364. doi: 10.1111/j.1471-8286.2006.01678.x
- Prosser SW, deWaard JR, Miller SE, Hebert PD (2016) DNA barcodes from century-old type specimens using next-generation sequencing. *Molecular Ecology Resources* 16: 487–497. doi: 10.1111/1755-0998.12474
- Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PLoS ONE* 8: e66213.
- Rubinoff D, Cameron S, Will K (2006) A genomic perspective on the shortcomings of mitochondrial DNA for “barcoding” identification. *Journal of Heredity* 97: 581–594. doi:1093/jhered/esl036
- Schmidt S, Schmid-Egger C, Moriniere J, Haszprunar G, Hebert PD (2015) DNA barcoding largely supports 250 years of classical taxonomy: identifications for Central European bees (Hymenoptera, *Apoidea partim*). *Molecular Ecology Resources* 15: 985–1000. doi: 10.1111/1755-0998.12363
- Serville A (1831) *Revue méthodique des insectes de l’ordre des Orthoptères*. Annales des sciences naturelles, Paris, 145 pp.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882. doi: 10.1093/nar/25.24.4876
- Versteirt V, Nagy ZT, Roelants P, Denis L, Breman FC, Damiens D, Dekoninck W, Backeljau T, Coosemans M, Van Bortel W (2015) Identification of Belgian mosquito species (Diptera: Culicidae) by DNA barcoding. *Molecular Ecology Resources* 15: 449–457. doi:10.1111/1755-0998.12318
- Virgilio M, Backeljau T, Nevado B, De Meyer M (2010) Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics* 11: 206. doi: 10.1186/1471-2105-11-206
- Walker F (1869) *Catalogue of the specimens of Dermaptera Saltatoria*. British Museum, 393 pp.
- Wang JF, Shi FM (2005) A taxonomic review of the genus *Hexacentrus* Serville from China (Orthoptera: Conocephalidae). *Acta Entomologica Sinica* 48: 242–246.

- Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* 54: 844–851. doi: 10.1080/10635150500354878
- Zhou ZJ, Li RL, Huang DW, Shi FM (2012) Molecular identification supports most traditional morphological species of *Ruspolia* (Orthoptera: Conocephalinae). *Invertebrate Systematics* 26: 451–456. doi: 10.1071/IS12019
- Zhou ZJ, Shang N, Chang YL, Shi FM (2014) DNA barcodes reveal intra- and inter- individual sequence variation in the bush cricket *Mecopoda niponensis* (Orthoptera: Tettigoniidae). *Biotechnology Bulletin* 5: 129–136.